

FORM PTO 1390
(REV 5-93)

US DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

ATTORNEY DOCKET NUMBER
2001_1023ATRANSMITTAL LETTER TO THE UNITED STATES
DESIGNATED/ELECTED OFFICE (DO/EO/US)
CONCERNING A FILING UNDER 35 U.S.C. §371U.S. APPLICATION NO.
(if known, see 37 CFR 1.55)
NEW 09/889722International Application No.
PCT/JP00/08253International Filing Date
November 22, 2000Priority Date Claimed
November 24, 1999**Title of Invention**

A HUMAN NUCLEAR PROTEIN HAVING A WW DOMAIN AND A POLYNUCLEOTIDE ENCODING THE PROTEIN

Applicant(s) For DO/EO/US

Seishi KATO, Akihiko KOMURO, Yutaka HIROSE

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. §371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. §371.
3. ☐ This express request to begin national examination procedures (35 U.S.C. §371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. §371(b) and PCT Articles 22 and 39(1).
4. ☐ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. §371(c)(2))
 - a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☒ has been transmitted by the International Bureau.
 - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)
6. ☒ A translation of the International Application into English (35 U.S.C. §371(c)(2)). **ATTACHMENT A**
7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. §371(c)(3)).
 - a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☐ have been transmitted by the International Bureau.
 - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
 - d. ☐ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19.
9. ☒ An **unexecuted** executed oath or declaration of the inventor(s) (35 U.S.C. §371(c)(4)). **ATTACHMENT B**
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. §371(c)(5)).

Items 11. to 14. below concern other document(s) or information included:

11. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98. **ATTACHMENT C**
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☒ A **FIRST** preliminary amendment. **ATTACHMENT D**
 - ☐ A **SECOND** or **SUBSEQUENT** preliminary amendment.
14. ☒ Other items or information: DISKETTE CONTAINING SEQUENCE LISTING

THE COMMISSIONER IS AUTHORIZED
TO CHARGE ANY DEFICIENCY IN THE
FEE FOR THIS PAPER TO DEPOSIT
ACCOUNT NO. 23-0975.

U.S. APPLICATION NO. (of U.S. Patent)
NEW

09/889722

INTERNATIONAL APPLICATION NO.
PCT/JP00/08253ATTORNEY'S DOCKET NO.
2001 1023A15. ☒ The following fees are submitted**BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):**

Neither international preliminary examination fee nor international search fee paid to USPTO
and International Search Report not prepared by the EPO or JPO \$1000.00
International Search Report has been prepared by the EPO or JPO \$ 860.00
International preliminary examination fee not paid to USPTO but international search
paid to USPTO \$ 710.00
International preliminary examination fee paid to USPTO but claims did not satisfy provisions
of PCT Article 33(1)-(4) \$ 690.00
International preliminary examination fee paid to USPTO and all claims satisfied provisions of
PCT Article 33(1)-(4) \$ 100.00

ENTER APPROPRIATE BASIC FEE AMOUNT =

\$860.00

Surcharge of \$130.00 for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest
claimed priority date (37 CFR 1.492(e)).

\$

Claims	Number Filed	Number Extra	Rate		
Total Claims	-20 =		X \$18.00	\$	
Independent Claims	- 3 =		X \$80.00	\$	
Multiple dependent claim(s) (if applicable)			+ \$270.00	\$	

TOTAL OF ABOVE CALCULATIONS =

\$860.00

☐ Small Entity Status is hereby asserted. Above fees are reduced by 1/2.

\$

SUBTOTAL =

\$860.00

Processing fee of \$130.00 for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest
claimed priority date (37 CFR 1.492(f)).

+

\$

TOTAL NATIONAL FEE =

\$860.00

Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an
appropriate cover sheet (37 CFR 3.28, 3.31). \$40 per property +

\$

TOTAL FEES ENCLOSED =

\$860.00

Amount to be refunded \$

Amount to be charged \$

a. ☒ A check in the amount of \$860.00 to cover the above fees is enclosed. A duplicate copy of this form is enclosed.b. ☐ Please charge my Deposit Account No. 23-0975 in the amount of \$_____ to cover the above fees.
A duplicate copy of this sheet is enclosed.c. ☐ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any
overpayment to Deposit Account No. 23-0975.**NOTE:** Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or
(b)) must be filed and granted to restore the application to pending status.

19. CORRESPONDENCE ADDRESS



000513

PATENT TRADEMARK OFFICE

By:

Warren M. Cheek, Jr.,
Registration No. 33,367WENDEROTH, LIND & PONACK, L.L.P.
2033 "K" Street, N.W., Suite 800
Washington, D.C. 20006-1021
Phone: (202) 721-8200
Fax: (202) 721-8250

July 20, 2001

[CHECK NO. 45524]

[2001 1023A]

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of :
Seishi KATO et al. : Attn: BOX PCT
Serial No. NEW : Docket No. 2001_1023A
Filed July 20, 2001 :

A HUMAN NUCLEAR PROTEIN HAVING A WW
DOMAIN AND A POLYNUCLEOTIDE ENCODING
THE PROTEIN
[Corresponding to PCT/JP00/08253
Filed November 22, 2000]

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents,
Washington, DC 20231

Sir:

Prior to calculating the filing fee, please amend the above-identified application as follows:

IN THE SPECIFICATION

Page 1, immediately after the title, please insert:

This application is a 371 of PCT/JP00/08253 filed November 22, 2000.

IN THE CLAIMS

Please amend the claims as follows:

5. (Amended) An expression vector expressing the polynucleotide of claim 2 in *in vitro* translation or in host cells.

ATTACHMENT D

protein of claim 1, and which polynucleotide comprises the nucleotide sequence of SEQ ID NO.

2.

Please add the following new claims:

8. An expression vector expressing the polynucleotide of claim 3 in *in vitro* translation or in host cells.

9. A transformed cell producing the human nuclear protein of claim 1, which is a cell transformed with an expression vector which expresses a polynucleotide encoding the protein of claim 1, and which polynucleotide consists of the nucleotide sequence of SEQ ID NO. 2.

REMARKS


The specification has been amended to reflect the 371 status. In addition, the claims have been amended to remove the multiple dependencies to reduce the PTO filing fee.

Attached hereto is a marked-up version of the changes made to the specification and claims by the current amendment. The attached pages are captioned "**Version with markings to show changes made**".

Favorable action on the merits is solicited.

Respectfully submitted,

Seishi KATO et al.

By 
Warren M. Cheek, Jr.
Registration No. 33,367
Attorney for Applicants

WMC/dlk
Washington, D.C. 20006-1021
Telephone (202) 721-8200
Facsimile (202) 721-8250
July 20, 2001

CLAIMS

1. An isolated and purified human nuclear protein comprising the amino acid sequence of SEQ ID NO: 1.

5

2. A polynucleotide encoding the protein of claim 1, which comprises the nucleotide sequence of SEQ ID NO: 2.

3. The polynucleotide of claim 2, consisting of the nucleotide sequence of SEQ ID NO: 2.

10

4. A human genomic DNA fragment with which a polynucleotide of SEQ ID NO:3 or a partial contiguous sequence thereof hybridizes under stringent conditions.

15

5. (Amended) An expression vector expressing the polynucleotide of claim 2 ~~or 3~~ in *in vitro* translation or in host cells.

6. (Amended) A transformed cell producing the human nuclear protein of claim 1, which is ~~transformants with the expression vector of claim 5.~~

20

7. An antibody against the human nuclear protein of claim 1.

a cell transformed with an expression vector which expresses a polynucleotide encoding the protein of Claim 1, and which comprises the nucleotide sequence of SEQ. ID NO. 2.

polynucleotide

3

THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of

Seishi KATO et al.

Serial No. 09/889,722

Filed July 20, 2001

Attn: BOX PCT

Docket No. 2001_1023A

A HUMAN NUCLEAR PROTEIN HAVING A
WW DOMAIN AND A POLYNUCLEOTIDE
ENCODING THE PROTEIN

[Corresponding to PCT/JP00/08253

Filed November 22, 2000]

THE COMMISSIONER IS AUTHORIZED
TO CHARGE ANY DEFICIENCY IN THE
FEE FOR THIS PAPER TO DEPOSIT
ACCOUNT NO. 23-0075

AMENDMENT AND REPLY TO NOTIFICATION OF MISSING REQUIREMENTS
UNDER 35 USC 371

Assistant Commissioner for Patents,
Washington, DC 20231

Sir:

In response to the PTO Notification of Missing Requirements Under 35 USC 371 dated September 10, 2001, submitted herewith is a Declaration for the above application executed by the inventors.

Enclosed is a paper and computer readable copy of the Sequence Listing. Please replace the Sequence Listing originally filed with the attached substitute Sequence Listing. No new matter is added.

Also enclosed are the PTO surcharge of \$130.00 required by 37 CFR 1.492(e), and a copy of the PTO notice.

It is respectfully submitted that the application is now complete, and early indication thereof is now requested.

Respectfully submitted,
Seishi KATO et al.

By Warren M. Cheek, Jr.
Warren M. Cheek, Jr.
Registration No. 33,367
Attorney for Applicants

WMC/dlk
Washington, D.C. 20006-1021
Telephone (202) 721-8200
Facsimile (202) 721-8250
October 19, 2001

09889722 104901

DESCRIPTION**A Human Nuclear Protein having a WW Domain and
A Polynucleotide encoding the Protein**

5 *This application is a 371 of PCT/JP00/08253 filed November 22, 2000.*

Technical Field

10 The present invention relates to a novel protein having a WW domain
and existing in human cell nuclei, a polynucleotide encoding this protein, and
an antibody against this protein. The protein and antibody of the present
invention are useful for diagnosis and therapy of various diseases, and the
polynucleotide of the present invention is useful as a probe for genetic diagnosis
15 or as a genetic source for gene therapy. Further, the polynucleotide can be
used as a genetic source for large-scale production of the protein of this
invention.

Background Art

20 The term "nuclear protein" is a generic name of proteins functioning in
cell nucleus. In nucleus there are genomic DNA serving as a plan of organism,
and nuclear proteins are involved in replication, transcriptional regulation etc.
of these genomic DNA. Typical nuclear proteins whose functions have been
25 revealed include a transcription factor, a splicing factor, an intranuclear
receptor, a cell cycle regulator and a tumor suppressor. These factors are
closely related not only to life phenomena such as development and
differentiation but also to diseases such as cancers (New Medical Science,
"Tensha No Shikumi To Shikkan" (Mechanism of Transcription and Diseases) ed.
30 by Masahiro Muramatsu). Accordingly, these nuclear proteins are expected as

DESCRIPTION**A Human Nuclear Protein having a WW Domain and
A Polynucleotide encoding the Protein**

5

Technical Field

10 The present invention relates to a novel protein having a WW domain
and existing in human cell nuclei, a polynucleotide encoding this protein, and
an antibody against this protein. The protein and antibody of the present
invention are useful for diagnosis and therapy of various diseases, and the
polynucleotide of the present invention is useful as a probe for genetic diagnosis
15 or as a genetic source for gene therapy. Further, the polynucleotide can be
used as a genetic source for large-scale production of the protein of this
invention.

Background Art

20

The term "nuclear protein" is a generic name of proteins functioning in
cell nucleus. In nucleus there are genomic DNA serving as a plan of organism,
and nuclear proteins are involved in replication, transcriptional regulation etc.
of these genomic DNA. Typical nuclear proteins whose functions have been
25 revealed include a transcription factor, a splicing factor, an intranuclear
receptor, a cell cycle regulator and a tumor suppressor. These factors are
closely related not only to life phenomena such as development and
differentiation but also to diseases such as cancers (New Medical Science,
"Tensha No Shikumi To Shikkan" (Mechanism of Transcription and Diseases) ed.
30 by Masahiro Muramatsu). Accordingly, these nuclear proteins are expected as

target proteins for developing low-molecular pharmaceutical preparations that regulate transcription and translation of specific genes, and it is desired to obtain as many nuclear proteins as possible.

5 The WW domain belongs to a new family of protein-protein interaction motifs resembling SH2, SH3, PH and PTB domains. It is known that this domain consists of about 40 amino acid residues containing 2 conserved tryptophan residues, and like the SH3 domain, binds to a proline-rich amino acid sequence (H. I. Chen and M. Sudol., Proc. Natl. Sci. 92, 7819-7823, 1995).
10 As a result of X-ray crystallographic analysis of a WW domain/ligand conjugate, it was revealed that the three-dimensional structure of the WW domain is different from that of SH3 (M. J. Macias et al., Nature, 382, 646-649, 1996). Like other protein motifs, the WW domain is contained in the cytoskeleton system (P. Bork and M. Sudol TIBS, 19, 531-533, 1994), in proteins
15 participating in the signal transduction system (H. I. Chen and M. Sudol., Proc. Natl. Sci., 92, 7819-7823, 1995), in a ubiquitin-protein ligase in the protein degradation system (O. Staub et al., EMBO J., 15, 2371-2380, 1996) and in a transcription activator (P. Bork and M. Sudol, TIBS, 19, 531-533, 1994), and is
20 believed to play an important role in the intracellular signal transduction system.

25 The object of the present invention is to provide a novel protein present in human cell nucleus, a polynucleotide encoding this protein, and an antibody against this nuclear protein.

Disclosure of Invention

30 To achieve the object described above, the present application provides the following inventions (1) to (7):

(1) An isolated and purified human nuclear protein comprising the amino acid sequence of SEQ ID NO: 1.

5 (2) A polynucleotide encoding the protein of the invention (1), which comprises the nucleotide sequence of SEQ ID NO: 2.

10 (3) The polynucleotide of the invention (2), consisting of the nucleotide sequence of SEQ ID NO: 2.

(4) A human genomic DNA fragment with which a polynucleotide of SEQ ID NO:3 or a partial contiguous sequence thereof hybridizes under stringent conditions.

15 (5) An expression vector expressing the polynucleotide of the invention (2) or (3) in *in vitro* translation or in host cells.

(6) A transformed cell producing the human nuclear protein of the invention (1), which is transformant with the expression vector of the invention
20 (5).

(7) An antibody against the human nuclear protein of the invention (1).

25

Best Mode for Carrying Out the Invention

The protein of the invention (1) can be obtained by a method of isolation thereof from human organs, cell lines etc., by a method of preparing the peptide through chemical synthesis on the basis of the amino acid sequence set forth in
30 SEQ ID NO: 1 or by a method of production thereof by recombinant DNA

technique using the polynucleotide encoding the amino acid sequence of SEQ ID NO: 1, among which the method with recombinant DNA technique is preferably used. For example, a vector harboring the polynucleotide of the invention (2) or (3) is subjected to *in vitro* transcription to prepare RNA which is then used as a template in *in vitro* translation, whereby the protein can be expressed *in vitro*. Further, by integrating the polynucleotide in a conventional method into a suitable expression vector, the protein encoded by the polynucleotide can be expressed in a large amount in procaryotes such as *E. coli*, *Bacillus subtilis* etc. or eucaryotes such as yeasts, insect cells and mammalian cells.

To produce the protein of the invention (1) by expressing the DNA through *in vitro* translation, the polynucleotide of the invention (2) or (3) is integrated in a vector harboring an RNA polymerase promoter (the invention (5)) and added the vector to an *in vitro* translation system such as a rabbit reticulocyte lysate or a wheat germ extract containing an RNA polymerase compatible with said promoter, whereby the protein of the invention (1) can be produced *in vitro*. The RNA polymerase promoter includes e.g. T7, T3 and SP6. The vector harboring such RNA polymerase promoter includes e.g. pKA1, pCDM8, pT3/T7 18, pT7/3 19, and pBluescript II.

To produce the protein of the invention (1) by expressing the DNA in microorganisms such as *E. coli*, the polynucleotide of the invention (2) or (3) is integrated in an expression vector harboring an origin capable of replication in microorganisms, a promoter, a ribosome-binding site, a DNA cloning site, a terminator etc. to prepare the expression vector (the invention (5)) which is then used for transformation of host cells, and by culturing the resulting transformant (the invention (6)), the protein encoded by said polynucleotide can be produced in a large amount in the microorganism. If an initiation codon and a termination codon have been added respectively to sites upstream and

downstream from an arbitrary translated region in said expression vector, a protein fragment containing the arbitrary region can be obtained by expressing the DNA. Alternatively, it can also be expressed as a fusion protein with another protein. By cleaving this fusion protein with a suitable protease, the part of only the protein encoded by said polynucleotide can be obtained. The expression vector for *E. coli* includes e.g. pUC series vectors, pBluescript II, pET expression system vectors and pGEX expression system vectors.

To produce the protein of the invention (1) by expressing the DNA in eucaryotes, the translated region of the polynucleotide of the invention (2) or (3) is integrated in an eucaryotic expression vector harboring a promoter, a splicing region, a poly(A)-additional site etc. to prepare the expression vector (the invention (5)) which is then used for transforming eucaryotic cells (the invention (6)), whereby the protein of the invention (1) can be produced in the eucaryotic cells. The expression vector includes e.g. pKAI, pCDM8, pSVK3, pMSG, pSVL, pBK-CMV, pBK-RSV, EBV vector, pRS and pYES2. If vectors such as pIND/V5-His, pFLAG-CMV-2, pEGFP-N1 and pEGFP-C1 are used, the protein of the present invention can also be expressed as a fusion protein having various tags such as His tag, FLAG tag and GFP added thereto. As the eucaryotic cells, mammalian cultured cells such as simian renal cells COS7 and Chinese hamster ovary cells CHO, budding yeasts, fission yeasts, silkworm cells and *Xenopus* oocytes are generally used, but insofar as the protein of the invention (1) can be expressed, any eucaryotic cells can be used. For introducing the expression vector into eucaryotic cells, conventional methods such as the electroporation method, calcium phosphate method, liposome method and DEAE-dextran method can be used.

For isolating and purifying the protein of the invention (1) from a culture after expression of the desired protein in the procaryotic or eucaryotic cells, separation techniques known in the art can be used in combination.

Such techniques include e.g. treatment with a denaturant such as urea or a surfactant, sonication, enzymatic digestion, salting-out or solvent precipitation, dialysis, centrifugation, ultrafiltration, gel filtration, SDS-PAGE, isoelectric focusing, ion-exchange chromatography, hydrophobic chromatography, affinity chromatography and reverse phase chromatography.

The protein of the invention (1) encompasses peptide fragments (each consisting of 5 or more amino acid residues) containing any partial amino acid sequence from the SEQ ID NO: 1. Such a peptide fragment can be used as an antigen for preparing the antibody of the present invention. Further, the protein of the invention (1) encompasses fusion proteins with another arbitrary protein. For example, fusion proteins with glutathione-S-transferase (GST) or green fluorescent protein (GFP), described in the Examples, can be mentioned.

The polynucleotide (cDNA) of the invention (2) or (3) can be cloned from a cDNA library derived from e.g. human cells. The cDNA is synthesized using poly(A)⁺RNA as a template extracted from human cells. The human cells may be either cultured cells or cells excised by an operation etc. from the human body. The cDNA can be synthesized by any methods such as the Okayama-Berg method (Okayama, H. and Berg, P., Mol. Cell Biol., 2, 161-170, 1982) and the Gubler-Hoffman method (Gubler, U. and Hoffman, J. Gene, 25, 263-269, 1983), but for efficiently obtaining full-length clones, the Capping method (Kato, S. et al., Gene, 150, 243-250, 1994) described in the Examples is preferably used.

The polynucleotide of the invention (2) comprises the nucleotide sequence of SEQ ID NO: 2, and for example, the polynucleotide consisting of the nucleotide sequence of SEQ ID NO: 3 has a 2669-bp nucleotide sequence containing a 2115-bp open reading frame (ORF). This ORF encodes a protein consisting of 704 amino acid residues. The polynucleotide of the invention (3)

comprises the 2115-bp nucleotide sequence (SEQ ID NO:2) constituting this ORF. By expressing the cDNA of the invention (2) or (3) in *E. coli* or animal cultured cells, an about 80-kDa protein was obtained. This protein binds to a C-terminal domain of RNA polymerase II, so it is considered to participate in transcriptional regulation.

Since the protein of the invention (1) is expressed in any tissues, the same clone as the polynucleotide of the invention (2) or (3) can be easily obtained from a human cDNA library prepared from human cells by screening the library with an oligonucleotide probe synthesized on the basis of the nucleotide sequence of the polynucleotide set forth in SEQ ID NO: 2 or 3. Alternatively, the objective cDNA can also be synthesized by polymerase chain reaction (PCR) by use of such oligonucleotides as primers.

Generally, polymorphism of human genes occurs frequently due to individual variations. Accordingly, those polynucleotides where in SEQ ID NO: 2 or 3, one or more nucleotides have been added, deleted and/or substituted with other nucleotides fall under the scope of the invention (3) or (4).

Accordingly, those proteins where in SEQ ID NO: 1, one or more amino acids have been added, deleted and/or substituted with other amino acids as a result of such alterations to nucleotides also fall under the scope of the invention (1) insofar as they have the activity of a protein having the amino acid sequence of SEQ ID NO: 1.

The polynucleotide of the invention (2) or (3) encompasses DNA fragments (10 bp or more) containing any partial nucleotide sequence from the sequence of SEQ ID NO: 2 or 3. Further, DNA fragments consisting of a sense or antisense strand thereof fall under the scope of this invention. These DNA fragments can be used as probes for genetic diagnosis.

The invention (4) is concerned with a human genomic DNA fragment with which the polynucleotide of SEQ ID NO: 3 or a partial contiguous sequence thereof hybridizes under stringent conditions. As used herein, the stringent conditions are that enables specific and detectable binding between the polynucleotide of SEQ ID NO: 3 or a partial contiguous sequence thereof (30 bp or more) and chromosome-derived genomic DNA. The stringent conditions are defined in terms of salt concentration, organic solvent (e.g., formamide), temperature and other known conditions. That is, stringency is increased by a decrease in salt concentration, by an increase in organic solvent concentration, or by an increase in hybridization temperature. For example, the stringent salt concentration is usually about 750 mM or less NaCl and about 75 mM or less trisodium citrate, more preferably about 500 mM or less NaCl and about 50 mM or less trisodium citrate and most preferably about 250 mM or less NaCl and about 25 mM or less trisodium citrate. The stringent organic solvent concentration is about 35 % or more formamide, most preferably about 50 % or more formamide. The stringent temperature condition is about 30 °C or more, more preferably about 37 °C or more and most preferably about 42 °C or more. The other conditions include hybridization time, the concentration of a detergent (e.g. SDS), the presence or absence of carrier DNA, etc., and by combining these conditions, varying stringency can be established. Further, the conditions for washing after hybridization also affects stringency. The washing conditions are also defined in terms of salt concentration and temperature, and the stringency of washing is increased by a decrease in salt concentration or by an increase in temperature. For example, the stringent salt condition for washing is about 30 mM or less NaCl and about 3 mM or less trisodium citrate, most preferably about 15 mM or less NaCl and about 1.5 mM or less trisodium citrate. The stringent temperature condition for washing is about 25 °C or more, more preferably about 42 °C or more and most preferably about 68 °C or more. The genomic DNA fragment of the invention (4) can be

isolated for example by subjecting a genome library prepared from human chromosomal DNA to screening by the above stringent hybridization with said polynucleotide as a probe and subsequent washing.

5 The genomic DNA fragment of the invention (4) comprises expression-regulating regions (promoter/enhancer and suppressor sequences, etc.) for the region coding for the protein of the invention (1). These expression-regulating regions are useful as a material for screening a material regulating *in vivo* expression of the protein of the invention (1).

10 The antibody of the invention (7) can be obtained from serum in an animal immunized with the protein of the invention (1) as an antigen. The antigen used may be a peptide chemically synthesized on the basis of the amino acid sequence of SEQ ID NO: 1 or the protein expressed in the eucaryotic or
15 procaryotic cells. Alternatively, the antibody can be prepared by introducing the above-described expression vector for eucaryotic cells through an injection or a gene gun into animal muscles or skin and then collecting serum (e.g., an invention in JP-7-313187A). As the animal, a mouse, rat, rabbit, goat, chicken or the like is used. If a hybridoma is produced by fusing myeloma cells with B
20 cells collected from the spleen in the immunized animal, a monoclonal antibody against the protein of the invention (1) can be produced by the hybridoma.

Examples

25 The present invention will be described in more detail by reference to the Examples, which however are not intended to limit the scope of the present invention. Basic procedures for DNA recombination and enzymatic reaction were in accordance with those described in a literature (Molecular Cloning, A
30 Laboratory Manual, Cold Spring Harbor Laboratory, 1989). Unless otherwise

specified, the restriction enzymes and various modifying enzymes used were products of Takara Shuzo Co., Ltd. The buffer composition in each enzymatic reaction, as well as reaction conditions, was followed instructions attached to the kits. Synthesis of cDNA was conducted according to a literature (Kato, S. et al., Gene, 150, 243-250, 1994).

(i) cDNA cloning

As a result of large-scale determination of the nucleotide sequences of cDNA clones selected from a human full-length cDNA library (described in WO97/03190), clone HP03494 was obtained. This clone had a structure made of a 291-bp 5'-untranslated region, a 2115-bp ORF and a 263-bp 3'-untranslated region (SEQ ID NO: 3). The ORF encodes a protein consisting of 704 amino acid residues.

Using the amino acid sequence (SEQ ID NO: 1) of this protein, a protein database was searched, but none of known proteins had homology to this protein. Further examination of GenBank by using the nucleotide sequence of its cDNA indicated that some ESTs (e.g. Accession No. A1758365) have 90 % or more homology thereto, but they are partial sequences, so whether or not they code for the same protein as the protein of this invention cannot be judged.

Examination of motif sequences indicated that as shown in Table 1, the region of from the 43- to 78-positions has homology to WW domains. Tryptophan residues at the 49- and 72-positions and a proline residue at the 75-position are amino acid residues conserved in every known WW domain.

Table 1

Protein	Position	Amino Acid Sequence	Accession No.
Conserved Sequence		— W — G — Y Y — N — W — P —	
HP03494	43	ELVHAGWEKQWSRRENRPYYFNRFNQLWEMPVLGQHD	
Npw38	46	EGLPPSWYKVFDPSCGLPYYNADTDLVSWLSPHDPNSV	BAA76400
Yap_Human	171	VPLPAGWEMAKTSS. GQRYFLNHIDQTTTWQDPRKAMLS	P46937
Yap_Chick-1	169	VPLPPGWEMAKTPS. GQRYFLNHIDQTTTWQDPRKAMLS	P46936
Yap_Mouse-1	156	VPLPAGWEMAKTSS. GQRYFLNHNDQTTTWQDPRKAMLS	P46938
Ned4_Mouse-1	40	SPLPPGWEEERQDVL. GRYYVNHESRRTQWKRPSPDDL	P46935
Ned4_Human-1	218	SPLPPGWEEERQDIL. GRYYVNHESRRTQWKRPQDNL	P46934
Ned4_Mouse-2	196	SGLPPGWEEKQDDR. GRSYYVDHNSKTTTWSKPTMDDP	P46935
Ned4_Human-2	375	SGLPPGWEEKQDER. GRSYYVDHNSRTTWTKPTVQATV	P46934
Dmd_Human	3055	TSVOGPWERAISP. KVPYYINHETQTTQWDHPKMTLY	P11532
Dmd_Mouse	3048	TSVOGPWERAISP. KVPYYINHETQTTQWDHPKMTLY	P11531
FE65_Rat	42	SDLPAGWMRVQDTS. GTYYWHI. PTGTTQWEPPGRASPS	P46933
Msb1/Human	249	IVLPPNWKTDARPE. GKIIYYHVITRQTQWDPPTWESPG	
IQGA_Human	679	GDNSKWKVHWVKG. GYYYYHNLETQEGGWDEPPNFVN	P46940
FBP11-1_Mouse	1WTEHKSPD. GRYYNTETKQSTWEKPDCLKTP	U40747
FBP11-2_Mouse	36	LLSKCPWKTYKSDS. GKPYYSNQTKESEWAKP.....	U40747

(ii) Northern blotting

Multi tissue Northern Blot (Clontech) having human tissue poly(A)⁺RNA blotted thereon was used as an mRNA source. As the probe, an *EcoRI*-*NotI* fragment of full-length HP03494 cDNA, labeled with a radioisotope by a random primer labeling kit (Pharmacia), was used. The conditions for Northern blotting hybridization followed the protocol attached to the kit. An about 3-kb hybridization band was obtained from the heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testicle, ovary, small intestine, colon and peripheral blood, suggesting that this protein is a housekeeping one.

(iii) Protein synthesis by *in vitro* translation

A plasmid vector harboring the polynucleotide (cDNA) of this invention was used to perform *in vitro* transcription/translation by a TnT rabbit reticulocyte lysate kit (a product of Promega). The expression product was

labeled with a radioisotope by adding [³⁵S] methionine. Any reaction was conducted according to the protocol attached to the kit. 2 µg of the plasmid was reacted at 30 °C for 90 minutes in a 25 µl reaction solution containing 12.5 µl T_NT rabbit reticulocyte lysate, 0.5 µl buffer (attached to the kit), 2 µl amino acid mixture (not containing methionine), 2 µl (0.35 MBq/µl) of [³⁵S] methionine (Amersham), 0.5 µl of T7 RNA polymerase and 20 U of RNasin. Then, 2 µl SDS sampling buffer (125 mM Tris-HCl, pH 6.8, 120 mM 2-mercaptoethanol, 2 % SDS solution, 0.025 % bromophenol blue, 20 % glycerol) was added to 3 µl of the reaction solution, and the mixture was treated by heating at 95 °C for 3 minutes and subjected to SDS-polyacrylamide gel electrophoresis. By autoradiography, the molecular weight of the translated product was determined. As a result, the translation product, which had a molecular weight of 80 kDa almost similar to the molecular weight (80,618) deduced from the ORF, was formed.

(iv) Expression of GST fusion protein in *E. coli*

The translated region was amplified by PCR where pHP03494 was used as a template while a 26-mer sense primer (SEQ ID NO: 4) starting at a translation initiation codon and having an *Eco*RI recognition site added thereto and a 26-mer antisense primer (SEQ ID NO: 5) terminating at a termination codon having a *Sa*II recognition site added thereto were used respectively as primers. The PCR product was digested with restriction enzyme *Eco*RI and inserted into *Eco*RI site in vector pGEX-5X-1 (Pharmacia). After its nucleotide sequence was confirmed, the resulting plasmid was used for transforming *E. coli* BL21. The transformant was cultured at 37 °C for 5 hours in LB medium, and IPTG was added thereto at a final concentration of 0.4 mM, followed by culturing at 37 °C for 2.5 hours. The microorganism was separated by centrifugation and lysed in a lysing solution (50 mM Tris-HCl (pH 7.5), 1 mM EDTA-1 % Triton X-100, 0.2 % SDS, 0.2 mM PMSF), frozen once at -80 °C, thawed, and disrupted by sonication. After centrifugation at 1000 x g for 30

minutes, glutathione Sepharose 4B was added to the supernatant and incubated at 4 °C for 1 hour. After the beads were sufficiently washed, a fusion protein was eluted with an eluent (10 mM Tris-50 mM glutathione). As a result, a GST-HP03494 fusion protein having a molecular weight of about 110 kDa was obtained.

(v) Preparation of antibody

Domestic rabbits were immunized with the above fusion protein as the antigen to give antiserum. First, an antiserum fraction precipitating by 40 % saturation with ammonium sulfate was applied onto a GST affinity column to remove GST antibody. Then, the unadsorbed fraction was purified by a GST-HP03494-antigen column.

(vi) Western blotting

A lysate of human fibrosarcoma cell line HT-1080 was separated by SDS-PAGE, blotted onto a PVDF membrane, blocked for 1 hour at room temperature with 0.05 % Tween 20-PBS (TPBS) containing 5 % skim milk, and incubated with the antibody diluted 10,000-fold with TPBS. The sample was washed 3 times with TPBS and then incubated for 1 hour with horseradish peroxidase-labeled goat anti-rabbit IgG diluted 10,000-fold with TPBS. The sample was washed four times with TPBS and detected by luminescence with an ECL reagent (Amersham), to give a signal with a molecular weight of 80 kDa. This molecular weight agreed with the molecular weight of the *in vitro* translated protein product in the rabbit cell-free translation system.

(vii) Expression of GFP fusion protein

The translated region was amplified by PCR where pHP03494 was used as a template while a 26-mer sense primer (SEQ ID NO: 4) starting at a translation initiation codon having an *EcoRI* recognition site added thereto and a 26-mer antisense primer (SEQ ID NO: 5) terminating at a termination codon

having a *SaII* recognition site added thereto were used respectively as primers. The PCR product was digested with restriction enzymes *EcoRI* and *SaII* and inserted into *EcoRI* site in GFP fusion protein expression vector pEGFP-C2 (Clontech). After the nucleotide sequence was confirmed, HeLa cells were transfected by the lipofection method with the resulting plasmid pEGFP-C2-HP03494. Under a fluorescence microscope, the cells transfected with pEGFP-C2 showed fluorescence on the whole of the cells, whereas the cells transfected with pEGFP-C2-HP03494 showed fluorescence on their nuclei only. This result indicated that HP03494 is a protein present in nucleus.

(viii) Binding to a C-terminal domain (CTD) of RNA polymerase II

The translated region coding for WW domain was amplified by PCR where pHP03494 was used as a template while a 33-mer sense primer (SEQ ID NO: 6) starting at a translation initiation codon with a *BamHI* recognition site added thereto and a 33-mer antisense primer (SEQ ID NO: 7) terminating at a termination codon with an *EcoRI* recognition site added thereto were used respectively as primers. The PCR product was digested with restriction enzymes *BamHI* and *EcoRI* and then inserted into *BamHI-EcoRI* sites in vector pGEX-5X-1 (Pharmacia). The resulting plasmid was subjected to expression in *E. coli* in the same manner as in (iv), to give a fusion protein GST-HP03494WW consisting of GST and HP03494 WW domain, and this fusion protein was separated by SDS-PAGE, then transferred onto a PVDF membrane, incubated with ³²P-labeled GST-CTD or ³²P-labeled GST-pCTD (GST-phosphorylated CTD) phosphorylated depending on a nuclear extract (Hirose, Y and Manley, J. L., Nature, 395, 93-96, 1998), and detected by the Far Western method (Kaelin, Jr. et al., Cell, 70, 351-364, 1992). It was revealed that the WW domain on HP03494 binds more strongly to phosphorylated CTD. This result suggested that the protein of this invention is involved in regulating transcription.

Industrial Applicability

This invention provides an isolated and purified human nuclear protein existing in human cell nucleus, a polynucleotide (human cDNA and genomic DNA fragment) encoding this protein, and an antibody against this nuclear protein. The protein and antibody of this invention are useful for diagnosis and therapy of morbid states such as cancers. By use of the present polynucleotide, the present protein can be expressed in a large amount. By screening a low-molecular compound binding to the present protein, a new type of pharmaceutical preparation such as antitumor agent can be searched for.

09889722-101901

CLAIMS

1. An isolated and purified human nuclear protein comprising the amino acid sequence of SEQ ID NO: 1.

5

2. A polynucleotide encoding the protein of claim 1, which comprises the nucleotide sequence of SEQ ID NO: 2.

10

3. The polynucleotide of claim 2, consisting of the nucleotide sequence of SEQ ID NO: 2.

15

4. A human genomic DNA fragment with which a polynucleotide of SEQ ID NO:3 or a partial contiguous sequence thereof hybridizes under stringent conditions.

20

5. An expression vector expressing the polynucleotide of claim 2 or 3 in *in vitro* translation or in host cells.

6. A transformed cell producing the human nuclear protein of claim 1, which is transformants with the expression vector of claim 5.

7. An antibody against the human nuclear protein of claim 1.

DECLARATION AND POWER OF ATTORNEY FOR U.S. PATENT APPLICATION

(X) Original () Supplemental () Substitute (X) PCT () DESIGN

As a below named inventor, I hereby declare that: my residence, post office address and citizenship are as stated below next to my name; that I verily believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural inventors are named below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

Title: A HUMAN NUCLEAR PROTEIN HAVING A WW DOMAIN AND A POLYNUCLEOTIDE ENCODING THE PROTEIN

of which is described and claimed in:

() the attached specification, or

(X) the specification in application Serial No. _____, filed July 20, 2001, and with amendments through _____, or

(X) the specification in International Application No. PCT/JP00/08253, filed November 22, 2000, and as amended on _____ (if applicable).

I hereby state that I have reviewed and understand the content of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

I acknowledge my duty to disclose to the Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby claim priority benefits under Title 35, United States Code, §119 (and §172 if this application is for a Design) of any application(s) for patent or inventor's certificate listed below and have also identified below any application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

COUNTRY	APPLICATION NO.	DATE OF FILING	PRIORITY CLAIMED
Japan	1999-332572	November 24, 1999	Yes

I hereby claim the benefit under Title 35, United States Code §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose information material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

APPLICATION SERIAL NO.	U.S. FILING DATE	STATUS: PATENTED, PENDING, ABANDONED

And I hereby appoint Michael R. Davis, Reg. No. 25,134; Matthew M. Jacob, Reg. No. 25,154; Warren M. Cheek, Jr., Reg. No. 33,367; Nils Pedersen, Reg. No. 33,145; Charles R. Watts, Reg. No. 33,142; and Michael S. Huppert, Reg. No. 40,268, who together constitute the firm of WENDEROTH, LIND & PONACK, L.L.P., as well as any other attorneys and agents associated with Customer No. 000513, to prosecute this application and to transact all business in the U.S. Patent and Trademark Office connected therewith.

I hereby authorize the U.S. attorneys and agents named herein to accept and follow instructions from NISHIZAWA & ASSOCIATES as to any action to be taken in the U.S. Patent and Trademark Office regarding this application without direct communication between the U.S. attorneys and myself. In the event of a change in the persons from whom instructions may be taken, the U.S. attorneys named herein will be so notified by me.

ATTACHMENT B

Direct Correspondence to Customer No:



000513

PATENT TRADEMARK OFFICE

Direct Telephone Calls to:

WENDEROTH, LIND & PONACK, L.L.P.
2033 "K" Street, N.W., Suite 800
Washington, D.C. 20006-1021

Phone: (202) 721-8200
Fax: (202) 721-8250

Full Name of First Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
	KATO	Seishi	
Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
	Kanagawa	Japan	Japan
Post Office Address	ADDRESS	CITY	STATE OR COUNTRY ZIP CODE
	46-50, Wakamatsu 3-chome, Sagamihara-shi, Kanagawa, Japan		
Full Name of Second Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
	KOMURO	Akihiko	
Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
	Kanagawa	Japan	Japan
Post Office Address	ADDRESS	CITY	STATE OR COUNTRY ZIP CODE
	2759-2, Kamitsuruma, Sagamihara-shi, Kanagawa, Japan		
Full Name of Third Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
	HIROSE	Yutaka	
Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
	Ishikawa	Japan	Japan
Post Office Address	ADDRESS	CITY	STATE OR COUNTRY ZIP CODE
	D-8 Wakunami-shukusha, 7-10, Wakunami 2-chome, Kanazawa-shi, Ishikawa, Japan		
Full Name of Fourth Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
Post Office Address	ADDRESS	CITY	STATE OR COUNTRY ZIP CODE
Full Name of Fifth Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
Post Office Address	ADDRESS	CITY	STATE OR COUNTRY ZIP CODE
Full Name of Sixth Inventor	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME

Residence & Citizenship	CITY	STATE OR COUNTRY	COUNTRY OF CITIZENSHIP
Post Office Address	ADDRESS	CITY	STATE OR COUNTRY ZIP CODE

I further declare that all statements made herein of my own knowledge are true, and that all statements on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

1st Inventor Seishi Kato Date August 28, 2001
 Seishi KATO
 2nd Inventor Akihiko Komuro Date August 28, 2001
 Akihiko KOMURO
 3rd Inventor Yutaka Hirose Date August 28, 2001
 Yutaka HIROSE
 4th Inventor _____ Date _____
 5th Inventor _____ Date _____
 6th Inventor _____ Date _____

The above application may be more particularly identified as follows:

U.S. Application Serial No. _____ Filing Date July 20, 2001

Applicant Reference Number 00-F-061PCT-US/YS Atty Docket No. 2001 1023A

Title of Invention A HUMAN NUCLEAR PROTEIN HAVING A WW DOMAIN AND A POLYNUCLEOTIDE ENCODING THE PROTEIN

SEQUENCE LISTING

<110> Japan Science and Technology Corporation

<120> Human nucleoprotein having a WW domain and
a polynucleotide encoding the protein

<130> 00-F-061PCT

<140> PCT/JP00/08253

<141> 2000-11-22

<150> JP11-332572

<151> 1999-11-24

<160> 7

<170> PatentIn Ver. 2.0

<210> 1

<211> 704

<212> PRT

<213> Homo sapiens

<400> 1

Met Ala Asn Glu Asn His Gly Ser Pro Arg Glu Glu Ala Ser Leu Leu

1

5

10

15

Ser His Ser Pro Gly Thr Ser Asn Gln Ser Gln Pro Cys Ser Pro Lys

20

25

30

Pro Ile Arg Leu Val Gln Asp Leu Pro Glu Glu Leu Val His Ala Gly

0989722.101901

35	40	45
Trp Glu Lys Cys Trp Ser Arg Arg Glu Asn Arg Pro Tyr Tyr Phe Asn		
50	55	60
Arg Phe Thr Asn Gln Ser Leu Trp Glu Met Pro Val Leu Gly Gln His		
65	70	75
Asp Val Ile Ser Asp Pro Leu Gly Leu Asn Ala Thr Pro Leu Pro Gln		
85	90	95
Asp Ser Ser Leu Val Glu Thr Pro Pro Ala Glu Asn Lys Pro Arg Lys		
100	105	110
Arg Gln Leu Ser Glu Glu Gln Pro Ser Gly Asn Gly Val Lys Lys Pro		
115	120	125
Lys Ile Glu Ile Pro Val Thr Pro Thr Gly Gln Ser Val Pro Ser Ser		
130	135	140
Pro Ser Ile Pro Gly Thr Pro Thr Leu Lys Met Trp Gly Thr Ser Pro		
145	150	155
Glu Asp Lys Gln Gln Ala Ala Leu Leu Arg Pro Thr Glu Val Tyr Trp		
165	170	175
Asp Leu Asp Ile Gln Thr Asn Ala Val Ile Lys His Arg Gly Pro Ser		
180	185	190
Glu Val Leu Pro Pro His Pro Glu Val Glu Leu Leu Arg Ser Gln Leu		
195	200	205
Ile Leu Lys Leu Arg Gln His Tyr Arg Glu Leu Cys Gln Gln Arg Glu		
210	215	220
Gly Ile Glu Pro Pro Arg Glu Ser Phe Asn Arg Trp Met Leu Glu Arg		
225	230	235
Lys Val Val Asp Lys Gly Ser Asp Pro Leu Leu Pro Ser Asn Cys Glu		
245	250	255
Pro Val Val Ser Pro Ser Met Phe Arg Glu Ile Met Asn Asp Ile Pro		
260	265	270

09669722 101601

Ile	Arg	Leu	Ser	Arg	Ile	Lys	Phe	Arg	Glu	Glu	Ala	Lys	Arg	Leu	Leu
275							280					285			
Phe	Lys	Tyr	Ala	Glu	Ala	Ala	Arg	Arg	Leu	Ile	Glu	Ser	Arg	Ser	Ala
290							295					300			
Ser	Pro	Asp	Ser	Arg	Lys	Val	Val	Lys	Trp	Asn	Val	Glu	Asp	Thr	Phe
305						310					315				320
Ser	Trp	Leu	Arg	Lys	Asp	His	Ser	Ala	Ser	Lys	Glu	Asp	Tyr	Met	Asp
						325					330				335
Arg	Leu	Glu	His	Leu	Arg	Arg	Gln	Cys	Gly	Pro	His	Val	Ser	Ala	Ala
						340					345				350
Ala	Lys	Asp	Ser	Val	Glu	Gly	Ile	Cys	Ser	Lys	Ile	Tyr	His	Ile	Ser
						355					360				365
Leu	Glu	Tyr	Val	Lys	Arg	Ile	Arg	Glu	Lys	His	Leu	Ala	Ile	Leu	Lys
						370					375				380
Glu	Asn	Asn	Ile	Ser	Glu	Glu	Val	Glu	Ala	Pro	Glu	Val	Glu	Pro	Arg
385						390					395				400
Leu	Val	Tyr	Cys	Tyr	Pro	Val	Arg	Leu	Ala	Val	Ser	Ala	Pro	Pro	Met
						405					410				415
Pro	Ser	Val	Glu	Met	His	Met	Glu	Asn	Asn	Val	Val	Cys	Ile	Arg	Tyr
						420					425				430
Lys	Gly	Glu	Met	Val	Lys	Val	Ser	Arg	Asn	Tyr	Phe	Ser	Lys	Leu	Trp
						435					440				445
Leu	Leu	Tyr	Arg	Tyr	Ser	Cys	Ile	Asp	Asp	Ser	Ala	Phe	Glu	Arg	Phe
						450					455				460
Leu	Pro	Arg	Val	Trp	Cys	Leu	Leu	Arg	Arg	Tyr	Gln	Met	Met	Phe	Gly
465						470					475				480
Val	Gly	Leu	Tyr	Glu	Gly	Thr	Gly	Leu	Gln	Gly	Ser	Leu	Pro	Val	His
						485					490				495
Val	Phe	Glu	Ala	Leu	His	Arg	Leu	Phe	Gly	Val	Ser	Phe	Glu	Cys	Phe

500 505 510
 Ala Ser Pro Leu Asn Cys Tyr Phe Arg Gln Tyr Cys Ser Ala Phe Pro
 515 520 525
 Asp Thr Asp Gly Tyr Phe Gly Ser Arg Gly Pro Cys Leu Asp Phe Ala
 530 535 540
 Pro Leu Ser Gly Ser Phe Glu Ala Asn Pro Pro Phe Cys Glu Glu Leu
 545 550 555 560
 Met Asp Ala Met Val Ser His Phe Glu Arg Leu Leu Glu Ser Ser Pro
 565 570 575
 Glu Pro Leu Ser Phe Ile Val Phe Ile Pro Glu Trp Arg Glu Pro Pro
 580 585 590
 Thr Pro Ala Leu Thr Arg Met Glu Gln Ser Arg Phe Lys Arg His Gln
 595 600 605
 Leu Ile Leu Pro Ala Phe Glu His Glu Tyr Arg Ser Gly Ser Gln His
 610 615 620
 Ile Cys Lys Lys Glu Glu Met His Tyr Lys Ala Val His Asn Thr Ala
 625 630 635 640
 Val Leu Phe Leu Gln Asn Asp Pro Gly Phe Ala Lys Trp Ala Pro Thr
 645 650 655
 Pro Glu Arg Leu Gln Glu Leu Ser Ala Ala Tyr Arg Gln Ser Gly Arg
 660 665 670
 Ser His Ser Ser Gly Ser Ser Ser Ser Ser Ser Ser Glu Ala Lys Asp
 675 680 685
 Arg Asp Ser Gly Arg Glu Gln Gly Pro Ser Arg Glu Pro His Pro Thr
 690 695 700

<210> 2

<211> 2112

00889722-101901

<213> Homo sapiens

atggccaatg agaatcacgg cagccccggg gaggaagcgt ccttgctgag tcactcccca 60
ggtacctcca atcagagcca gccctgttct ccaaagccaa tccgcctggt tcaggacctc 120
ccagaggagc tgggtcatgc aggctgggag aagtgtctgga gccggaggga gaatcgtccc 180
tactacttca accgattcac caaccagtcc ctgtgggaga tgcccgtgct ggggcagcac 240
gatgtgattt cggacccttt ggggctgaat gcgaccccac tgccccaaga ctcaagcttg 300
gtggaaactc ccccggtga gaacaagccc agaaagcggc agctctcgga agagcagcca 360
agcggcaatg gtgtgaagaa gcccagatt gaaatcccag tgacaccac aggccagtcg 420
gtgcccagct ccccagtat ccaggaacc ccaacgctga agatgtgggg tacgtcccct 480
gaagataaac agcaggcagc tctcctacga cccactgagg tctactggga cctggacatc 540
cagaccaatg ctgtcatcaa gcaccggggg ccttcagagg tgctgcccc gcatcccgaa 600
gtggaactgc tccgctctca gctcatcctg aagcttcggc agcactatcg ggagctgtgc 660
cagcagcgag agggcattga gcctccacgg gagtctttca accgctggat gctggagcgc 720
aaggtagtag acaaaggatc tgacccccctg ttgccagca actgtgaacc agtcgtgtca 780
ccttccatgt ttcgtgaaat catgaacgac attcctatca ggttatccc aatcaagttc 840
cgggaggaag ccaagcgctt gctctttaaa tatgcggagg ccgccaggcg gctcatcgag 900
tccaggagtg catccccga cagtaggaag gtggtcaa atggaatgtgga agacaccttt 960
agctggcttc ggaaggacca ctacgcctcc aaggaggact acatggatcg cctggagcat 1020
ctgcggaggc agtgtggccc ccacgtctcg gccgcagcca aggactccgt ggaaggcatc 1080
tgtagtaaga tctaccacat ctccctggag tacgtcaaac ggatccgaga gaagcacctt 1140
gccatcctca aggaaaacaa catctcagag gaggtggagg cccctgaggt ggagccccgc 1200
ctagtgtact gctaccagct ccggctggct gtgtctgcac cgcccatgcc cagcgtggag 1260
atgcacatgg agaacaacgt ggtctgcac cggataaagg gagagatggt caaggtcagc 1320
cgcaactact tcagcaagct gtggctcctt taccgctaca gctgcattga tgactctgcc 1380
tttgagaggt tcctgccccg ggtctggtgt cttctccgac ggtaccagat gatgttcggc 1440
gtgggcctct acgaggggac tggcctgcag ggatcgctgc ctgtgcatgt ctttgaggcc 1500

ctccaccgac tctttggcgt cagcttcgag tgcctcgcct caccctcaa ctgctacttc 1560
 cgccagtact gttctgcctt ccccgacaca gacggctact ttggctcccg cgggccctgc 1620
 ctagactttg ctccactgag tggttcattt gaggccaacc ctcccttctg cgaggagctc 1680
 atggatgcca tggctcttca ctttgagaga ctgcttgaga gctcaccgga gcccctgtcc 1740
 ttcatcgtgt tcatccctga gtggcgggaa cccccaacac cagcgtcac ccgcatggag 1800
 cagagccgct tcaaacgcca ccagttgatc ctgcctgcct ttgagcatga gtaccgcagt 1860
 ggctcccagc acatctgcaa gaaggaggaa atgcactaca aggccgtcca caacacggct 1920
 gtgctcttcc tacagaacga ccctggcttt gccaaagtgg cgccgacgcc tgaacggctg 1980
 caggagctga gtgtgccta ccggcagtcg ggccgcagcc acagctctgg ttcttctca 2040
 tcgtctcct cggaggccaa ggaccgggac tcgggccgtg agcagggtcc tagccgcgag 2100
 cctcacccca ct 2112

<210> 3

<211> 2669

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (292).. (2406)

<400> 3

acacaagatg gcggcagcgg cgctggggag ggcgaggcgg aggcggcaaa acgggcggtc 60
 gagcagaacg tgtagccgcg tccctccag tccgtccgg gcagctgctg atgcaaggaa 120
 tccctgggc tccgtccac tccactgctg accagcccat tcgctgtgc tgagtcttcc 180
 tgcaggcctt tccttgctc tgtgggaccc tgtgggggtc catccggctg gagaagaaaa 240
 gcctctcatg ctaacgttgc agacccaga gggctctgtg tgggtgtgga g atg gcc 297

Met Ala

0989722.101904

aat gag aat cac ggc agc ccc cgg gag gaa gcg tcc ctg ctg agt cac 345
 Asn Glu Asn His Gly Ser Pro Arg Glu Glu Ala Ser Leu Leu Ser His
 5 10 15
 tcc cca ggt acc tcc aat cag agc cag ccc tgt tct cca aag cca atc 393
 Ser Pro Gly Thr Ser Asn Gln Ser Gln Pro Cys Ser Pro Lys Pro Ile
 20 25 30
 cgc ctg gtt cag gac ctc cca gag gag ctg gtg cat gca ggc tgg gag 441
 Arg Leu Val Gln Asp Leu Pro Glu Glu Leu Val His Ala Gly Trp Glu
 35 40 45 50
 aag tgc tgg agc cgg agg gag aat cgt ccc tac tac ttc aac cga ttc 489
 Lys Cys Trp Ser Arg Arg Glu Asn Arg Pro Tyr Tyr Phe Asn Arg Phe
 55 60 65
 acc aac cag tcc ctg tgg gag atg ccc gtg ctg ggg cag cac gat gtg 537
 Thr Asn Gln Ser Leu Trp Glu Met Pro Val Leu Gly Gln His Asp Val
 70 75 80
 att tcg gac cct ttg ggg ctg aat gcg acc cca ctg ccc caa gac tca 585
 Ile Ser Asp Pro Leu Gly Leu Asn Ala Thr Pro Leu Pro Gln Asp Ser
 85 90 95
 agc ttg gtg gaa act ccc ccg gct gag aac aag ccc aga aag cgg cag 633
 Ser Leu Val Glu Thr Pro Pro Ala Glu Asn Lys Pro Arg Lys Arg Gln
 100 105 110
 ctc tcg gaa gag cag cca agc ggc aat ggt gtg aag aag ccc aag att 681
 Leu Ser Glu Glu Gln Pro Ser Gly Asn Gly Val Lys Lys Pro Lys Ile
 115 120 125 130
 gaa atc cca gtg aca ccc aca ggc cag tcg gtg ccc agc tcc ccc agt 729
 Glu Ile Pro Val Thr Pro Thr Gly Gln Ser Val Pro Ser Ser Pro Ser
 135 140 145
 atc cca gga acc cca acg ctg aag atg tgg ggt acg tcc cct gaa gat 777

09882210404
 105101 3328850

Ile Pro Gly Thr Pro Thr Leu Lys Met Trp Gly Thr Ser Pro Glu Asp	
150	155
160	
aaa cag cag gca gct ctc cta cga ccc act gag gtc tac tgg gac ctg	825
Lys Gln Gln Ala Ala Leu Leu Arg Pro Thr Glu Val Tyr Trp Asp Leu	
165	170
175	
gac atc cag acc aat gct gtc atc aag cac cgg ggg cct tca gag gtg	873
Asp Ile Gln Thr Asn Ala Val Ile Lys His Arg Gly Pro Ser Glu Val	
180	185
190	
ctg ccc ccg cat ccc gaa gtg gaa ctg ctc cgc tct cag ctc atc ctg	921
Leu Pro Pro His Pro Glu Val Glu Leu Leu Arg Ser Gln Leu Ile Leu	
195	200
205	210
aag ctt cgg cag cac tat cgg gag ctg tgc cag cag cga gag ggc att	969
Lys Leu Arg Gln His Tyr Arg Glu Leu Cys Gln Gln Arg Glu Gly Ile	
215	220
225	
gag cct cca cgg gag tct ttc aac cgc tgg atg ctg gag cgc aag gtg	1017
Glu Pro Pro Arg Glu Ser Phe Asn Arg Trp Met Leu Glu Arg Lys Val	
230	235
240	
gta gac aaa gga tct gac ccc ctg ttg ccc agc aac tgt gaa cca gtc	1065
Val Asp Lys Gly Ser Asp Pro Leu Leu Pro Ser Asn Cys Glu Pro Val	
245	250
255	
gtg tca cct tcc atg ttt cgt gaa atc atg aac gac att cct atc agg	1113
Val Ser Pro Ser Met Phe Arg Glu Ile Met Asn Asp Ile Pro Ile Arg	
260	265
270	
tta tcc cga atc aag ttc cgg gag gaa gcc aag cgc ctg ctc ttt aaa	1161
Leu Ser Arg Ile Lys Phe Arg Glu Glu Ala Lys Arg Leu Leu Phe Lys	
275	280
285	290
tat gcg gag gcc gcc agg cgg ctc atc gag tcc agg agt gca tcc cct	1209
Tyr Ala Glu Ala Ala Arg Arg Leu Ile Glu Ser Arg Ser Ala Ser Pro	
295	300
305	

gac agt agg aag gtg gtc aaa tgg aat gtg gaa gac acc ttt agc tgg 1257
 Asp Ser Arg Lys Val Val Lys Trp Asn Val Glu Asp Thr Phe Ser Trp
 310 315 320
 ctt cgg aag gac cac tca gcc tcc aag gag gac tac atg gat cgc ctg 1305
 Leu Arg Lys Asp His Ser Ala Ser Lys Glu Asp Tyr Met Asp Arg Leu
 325 330 335
 gag cat ctg cgg agg cag tgt ggc ccc cac gtc tgc gcc gca gcc aag 1353
 Glu His Leu Arg Arg Gln Cys Gly Pro His Val Ser Ala Ala Ala Lys
 340 345 350
 gac tcc gtg gaa ggc atc tgc agt aag atc tac cac atc tcc ctg gag 1401
 Asp Ser Val Glu Gly Ile Cys Ser Lys Ile Tyr His Ile Ser Leu Glu
 355 360 365 370
 tac gtc aaa cgg atc cga gag aag cac ctt gcc atc ctc aag gaa aac 1449
 Tyr Val Lys Arg Ile Arg Glu Lys His Leu Ala Ile Leu Lys Glu Asn
 375 380 385
 aac atc tca gag gag gtg gag gcc cct gag gtg gag ccc cgc cta gtg 1497
 Asn Ile Ser Glu Glu Val Glu Ala Pro Glu Val Glu Pro Arg Leu Val
 390 395 400
 tac tgc tac cca gtc cgg ctg gct gtg tct gca ccg ccc atg ccc agc 1545
 Tyr Cys Tyr Pro Val Arg Leu Ala Val Ser Ala Pro Pro Met Pro Ser
 405 410 415
 gtg gag atg cac atg gag aac aac gtg gtc tgc atc cgg tat aag gga 1593
 Val Glu Met His Met Glu Asn Asn Val Val Cys Ile Arg Tyr Lys Gly
 420 425 430
 gag atg gtc aag gtc agc cgc aac tac ttc agc aag ctg tgg ctc ctt 1641
 Glu Met Val Lys Val Ser Arg Asn Tyr Phe Ser Lys Leu Trp Leu Leu
 435 440 445 450
 tac cgc tac agc tgc att gat gac tct gcc ttt gag agg ttc ctg ccc 1689
 Tyr Arg Tyr Ser Cys Ile Asp Asp Ser Ala Phe Glu Arg Phe Leu Pro

0988323 101001
 T06T07 2268857

455	460	465	
cgg gtc tgg tgt ctt ctc cga cgg tac cag atg atg ttc ggc gtg ggc	1737		
Arg Val Trp Cys Leu Leu Arg Arg Tyr Gln Met Met Phe Gly Val Gly			
470	475	480	
ctc tac gag ggg act ggc ctg cag gga tcg ctg cct gtg cat gtc ttt	1785		
Leu Tyr Glu Gly Thr Gly Leu Gln Gly Ser Leu Pro Val His Val Phe			
485	490	495	
gag gcc ctc cac cga ctc ttt ggc gtc agc ttc gag tgc ttc gcc tca	1833		
Glu Ala Leu His Arg Leu Phe Gly Val Ser Phe Glu Cys Phe Ala Ser			
500	505	510	
ccc ctc aac tgc tac ttc cgc cag tac tgt tct gcc ttc ccc gac aca	1881		
Pro Leu Asn Cys Tyr Phe Arg Gln Tyr Cys Ser Ala Phe Pro Asp Thr			
515	520	525	530
gac ggc tac ttt ggc tcc cgc ggg ccc tgc cta gac ttt gct cca ctg	1929		
Asp Gly Tyr Phe Gly Ser Arg Gly Pro Cys Leu Asp Phe Ala Pro Leu			
535	540	545	
agt ggt tca ttt gag gcc aac cct ccc ttc tgc gag gag ctc atg gat	1977		
Ser Gly Ser Phe Glu Ala Asn Pro Pro Phe Cys Glu Glu Leu Met Asp			
550	555	560	
gcc atg gtc tct cac ttt gag aga ctg ctt gag agc tca ccg gag ccc	2025		
Ala Met Val Ser His Phe Glu Arg Leu Leu Glu Ser Ser Pro Glu Pro			
565	570	575	
ctg tcc ttc atc gtg ttc atc cct gag tgg cgg gaa ccc cca aca cca	2073		
Leu Ser Phe Ile Val Phe Ile Pro Glu Trp Arg Glu Pro Pro Thr Pro			
580	585	590	
gcg ctc acc cgc atg gag cag agc cgc ttc aaa cgc cac cag ttg atc	2121		
Ala Leu Thr Arg Met Glu Gln Ser Arg Phe Lys Arg His Gln Leu Ile			
595	600	605	610
ctg cct gcc ttt gag cat gag tac cgc agt ggc tcc cag cac atc tgc	2169		

09069700 101901

Leu Pro Ala Phe Glu His Glu Tyr Arg Ser Gly Ser Gln His Ile Cys
 615 620 625
 aag aag gag gaa atg cac tac aag gcc gtc cac aac acg gct gtg ctc 2217
 Lys Lys Glu Glu Met His Tyr Lys Ala Val His Asn Thr Ala Val Leu
 630 635 640
 ttc cta cag aac gac cct ggc ttt gcc aag tgg gcg ccg acg cct gaa 2265
 Phe Leu Gln Asn Asp Pro Gly Phe Ala Lys Trp Ala Pro Thr Pro Glu
 645 650 655
 cgg ctg cag gag ctg agt gct gcc tac cgg cag tca ggc cgc agc cac 2313
 Arg Leu Gln Glu Leu Ser Ala Ala Tyr Arg Gln Ser Gly Arg Ser His
 660 665 670
 agc tct ggt tct tcc tca tcg tcc tcc tcg gag gcc aag gac cgg gac 2361
 Ser Ser Gly Ser Ser Ser Ser Ser Ser Ser Glu Ala Lys Asp Arg Asp
 675 680 685 690
 tcg ggc cgt gag cag ggt cct agc cgc gag cct cac ccc act taa 2406
 Ser Gly Arg Glu Gln Gly Pro Ser Arg Glu Pro His Pro Thr
 695 700 705
 catatcctgc ggggaggagg agccccaggg gtgctagtct ggactgctgg gactcgggcc 2466
 cctggggcct cagagggacc cgggctgcca ctgacatatg aagattatgg ttctgccagg 2526
 gctccccctc ctgcctgtcc ccaagtcctc acctcaaact ccctccaagt cccatgtata 2586
 taggtcctga tgccttccca accccgcccc tcacctgtt gccacctgtt ttcatttgta 2646
 aaaggaaata cagaaacccc ccc 2669

<210> 4

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<213> Synthesized oligonucleotide

<400> 4

ccgaattcat ggccaatgag aatcac

26

<210> 5

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<213> Synthesized oligonucleotide

<400> 5

ccgtcgactt aagtggggtg aggctc

26

<210> 6

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<213> Synthesized oligonucleotide

<400> 6

cgaggatccg ttcaggacct cccagaggacg cta

33

098972-10904

<210> 7

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<213> Synthesized oligonucleotide

<400> 7

cgagaattcc gaaatcacat cgtgctgccc cag

33

0989723-101901

SEQUENCE LISTING

<110> Japan Science and Technology Corporation

<120> Human nucleoprotein having a WW domain and
a polynucleotide encoding the protein

<130> 09/889,722

<140> PCT/JP00/08253

<141> 2000-11-22

<150> JP11-332572

<151> 1999-11-24

<160> 7

<170> PatentIn Ver. 2.0

<210> 1

<211> 704

<212> PRT

<213> Homo sapiens

<400> 1

```

Met Ala Asn Glu Asn His Gly Ser Pro Arg Glu Glu Ala Ser Leu Leu
  1          5          10          15
Ser His Ser Pro Gly Thr Ser Asn Gln Ser Gln Pro Cys Ser Pro Lys
          20          25          30
Pro Ile Arg Leu Val Gln Asp Leu Pro Glu Glu Leu Val His Ala Gly
          35          40          45
Trp Glu Lys Cys Trp Ser Arg Arg Glu Asn Arg Pro Tyr Tyr Phe Asn
          50          55          60
Arg Phe Thr Asn Gln Ser Leu Trp Glu Met Pro Val Leu Gly Gln His
          65          70          75          80
Asp Val Ile Ser Asp Pro Leu Gly Leu Asn Ala Thr Pro Leu Pro Gln
          85          90          95
Asp Ser Ser Leu Val Glu Thr Pro Pro Ala Glu Asn Lys Pro Arg Lys
          100          105          110
Arg Gln Leu Ser Glu Glu Gln Pro Ser Gly Asn Gly Val Lys Lys Pro
          115          120          125
Lys Ile Glu Ile Pro Val Thr Pro Thr Gly Gln Ser Val Pro Ser Ser
          130          135          140
Pro Ser Ile Pro Gly Thr Pro Thr Leu Lys Met Trp Gly Thr Ser Pro
          145          150          155          160
Glu Asp Lys Gln Gln Ala Ala Leu Leu Arg Pro Thr Glu Val Tyr Trp
          165          170          175
Asp Leu Asp Ile Gln Thr Asn Ala Val Ile Lys His Arg Gly Pro Ser
          180          185          190
Glu Val Leu Pro Pro His Pro Glu Val Glu Leu Leu Arg Ser Gln Leu
          195          200          205
Ile Leu Lys Leu Arg Gln His Tyr Arg Glu Leu Cys Gln Gln Arg Glu
          210          215          220
Gly Ile Glu Pro Pro Arg Glu Ser Phe Asn Arg Trp Met Leu Glu Arg
          225          230          235          240
Lys Val Val Asp Lys Gly Ser Asp Pro Leu Leu Pro Ser Asn Cys Glu

```

245 250 255
 Pro Val Val Ser Pro Ser Met Phe Arg Glu Ile Met Asn Asp Ile Pro
 260 265 270
 Ile Arg Leu Ser Arg Ile Lys Phe Arg Glu Glu Ala Lys Arg Leu Leu
 275 280 285
 Phe Lys Tyr Ala Glu Ala Ala Arg Arg Leu Ile Glu Ser Arg Ser Ala
 290 295 300
 Ser Pro Asp Ser Arg Lys Val Val Lys Trp Asn Val Glu Asp Thr Phe
 305 310 315 320
 Ser Trp Leu Arg Lys Asp His Ser Ala Ser Lys Glu Asp Tyr Met Asp
 325 330 335
 Arg Leu Glu His Leu Arg Arg Gln Cys Gly Pro His Val Ser Ala Ala
 340 345 350
 Ala Lys Asp Ser Val Glu Gly Ile Cys Ser Lys Ile Tyr His Ile Ser
 355 360 365
 Leu Glu Tyr Val Lys Arg Ile Arg Glu Lys His Leu Ala Ile Leu Lys
 370 375 380
 Glu Asn Asn Ile Ser Glu Glu Val Glu Ala Pro Glu Val Glu Pro Arg
 385 390 395 400
 Leu Val Tyr Cys Tyr Pro Val Arg Leu Ala Val Ser Ala Pro Pro Met
 405 410 415
 Pro Ser Val Glu Met His Met Glu Asn Asn Val Val Cys Ile Arg Tyr
 420 425 430
 Lys Gly Glu Met Val Lys Val Ser Arg Asn Tyr Phe Ser Lys Leu Trp
 435 440 445
 Leu Leu Tyr Arg Tyr Ser Cys Ile Asp Asp Ser Ala Phe Glu Arg Phe
 450 455 460
 Leu Pro Arg Val Trp Cys Leu Leu Arg Arg Tyr Gln Met Met Phe Gly
 465 470 475 480
 Val Gly Leu Tyr Glu Gly Thr Gly Leu Gln Gly Ser Leu Pro Val His
 485 490 495
 Val Phe Glu Ala Leu His Arg Leu Phe Gly Val Ser Phe Glu Cys Phe
 500 505 510
 Ala Ser Pro Leu Asn Cys Tyr Phe Arg Gln Tyr Cys Ser Ala Phe Pro
 515 520 525
 Asp Thr Asp Gly Tyr Phe Gly Ser Arg Gly Pro Cys Leu Asp Phe Ala
 530 535 540
 Pro Leu Ser Gly Ser Phe Glu Ala Asn Pro Pro Phe Cys Glu Glu Leu
 545 550 555 560
 Met Asp Ala Met Val Ser His Phe Glu Arg Leu Leu Glu Ser Ser Pro
 565 570 575
 Glu Pro Leu Ser Phe Ile Val Phe Ile Pro Glu Trp Arg Glu Pro Pro
 580 585 590
 Thr Pro Ala Leu Thr Arg Met Glu Gln Ser Arg Phe Lys Arg His Gln
 595 600 605
 Leu Ile Leu Pro Ala Phe Glu His Glu Tyr Arg Ser Gly Ser Gln His
 610 615 620
 Ile Cys Lys Lys Glu Glu Met His Tyr Lys Ala Val His Asn Thr Ala
 625 630 635 640
 Val Leu Phe Leu Gln Asn Asp Pro Gly Phe Ala Lys Trp Ala Pro Thr
 645 650 655
 Pro Glu Arg Leu Gln Glu Leu Ser Ala Ala Tyr Arg Gln Ser Gly Arg
 660 665 670
 Ser His Ser Ser Gly Ser Ser Ser Ser Ser Ser Glu Ala Lys Asp
 675 680 685
 Arg Asp Ser Gly Arg Glu Gln Gly Pro Ser Arg Glu Pro His Pro Thr
 690 695 700

<210> 2
 <211> 2112
 <212> DNA
 <213> Homo sapiens

<400> 2
 atggccaatg agaatcacgg cagcccccg gaggaagcgt ccctgctgag tcaactcccca 60
 ggtacctcca atcagagcca gccctgttct ccaaagccaa tccgcctggg tcaggacctc 120
 ccagaggagc tgggtgcatgc aggctgggag aagtgtctgga gccggaggga gaatcgtccc 180
 tactacttca accgattcac caaccagtcc ctgtgggaga tgcccgtgct ggggcagcac 240
 gatgtgattt cggacccttt ggggctgaat gcgacccac tgccccaaga ctcaagcttg 300
 gtggaaactc ccccggtctga gaacaagccc agaaaagcggc agctctcgga agagcagcca 360
 agcggcaatg gtgtgaagaa gcccaagatt gaaatcccag tgacacccac aggccagtgc 420
 gtgccagct cccccagtat cccaggaacc ccaacgctga agatgtgggg tacgtcccct 480
 gaagataaac agcaggcagc tctcctacga cccactgagg tctactggga cctggacatc 540
 cagaccaatg ctgtcatcaa gcaccggggg ccttcagagg tgctgcccc gcaccccga 600
 gtggaaactgc tccgctctca gctcatcctg aagcttcggc agcactatcg ggagctgtgc 660
 cagcagcgag agggcattga gcctccacgg gagtctttca accgctggat gctggagcgc 720
 aagggtggtag acaaaggatc tgacccccctg ttgccagca actgtgaacc agtcgtgtca 780
 ccttccatgt ttcgtgaaat catgaacgac attcctatca gggtatccc aatcaagttc 840
 cgggaggaag ccaagcgcct gctctttaa tatgcggagg ccgcccaggc gctcatcgag 900
 tccaggagtgc catcccctga cagtaggaag gtggtcaa atggaatgtgga agacaccttt 960
 agctggcttc ggaaggacca ctcagcctcc aaggaggact acatggatcg cctggagcat 1020
 ctgcggaggc agtgtggccc ccacgtctcg gccgcagcca aggactccgt ggaaggcatc 1080
 tgcagtaaga tctaccacat ctccctggag tacgtcaaac ggatccgaga gaagcacctt 1140
 gccatcctca aggaaaacaa catctcagag gaggtggagg cccctgaggt ggagccccgc 1200
 ctagtgtact gctacccagt ccggctgggt gtgtctgcac cgcccatgcc cagcgtggag 1260
 atgcacatgg agaacaacgt ggtctgcatc cgggtataagg gagagatggt caaggtcagc 1320
 cgcaactact tcagcaagct gtggctcctt taccgctaca gctgcattga tgactctgcc 1380
 tttgagaggt tccctgcccc ggtctggtgt cttctccgac ggtaccagat gatgttcggc 1440
 gtgggcctct acgaggggac tggcctgcag ggatcgctgc ctgtgcatgt ctttgaggcc 1500
 ctccaccgac tctttggcgt cagcttcgag tgcttcgct caccctcaa ctgctacttc 1560
 cgccagtact gttctgcctt ccccgacaca gacggctact ttggctccc cgggccctgc 1620
 ctgagctttg ctccactgag tggttcattt gaggccaacc ctcccttctg cgaggagctc 1680
 atggatgcca tgggtctctca ctttgagaga ctgcttgaga gctcaccgga gccctgtcc 1740
 ttcacgtgtg tcatccctga gtggcgggaa cccccaacac cagcgctcac ccgcatggag 1800
 cagagccgct tcaaacgcca ccagttgate ctgcctgcct ttgagcatga gtaccgcagt 1860
 ggctcccagc acatctgcaa gaaggaggaa atgcactaca aggcctgcca caacacggct 1920
 gtgctcttcc tacagaacga ccctggcttt gccaaagtgg cgccgacgcc tgaacggctg 1980
 caggagctga gtgtgccta ccggcagtc ggcgcagcc acagctctgg ttcttcctca 2040
 tcgtcctcct cggaggccaa ggaccgggac tcgggcccgt agcagggtcc tagccgcgag 2100
 cctcacccca ct 2112

<210> 3
 <211> 2669
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (292)..(2406)

<400> 3
 acacaagatg gcggcagcgg cgctggggag ggcgaggcgg aggcggcaaa acgggcggtc 60

gagcagaacg tgtagccgcg tcccctccag tccgctccgg gcagctgctg atgcaaggaa 120
 tcccctgggc tcccgtccac tccactgctg accagcccat tcgcctgtgc tgagtcttcc 180
 tgcaggcctt tccttgctc tgtgggaccc tgtgggggtc catccggctg gagaagaaaa 240
 gcctctcatg ctaacgttgc agaccccaga gggctcctgtg tgggtgtgga g atg gcc 297
 Met Ala
 1
 aat gag aat cac ggc agc ccc cgg gag gaa gcg tcc ctg ctg agt cac 345
 Asn Glu Asn His Gly Ser Pro Arg Glu Glu Ala Ser Leu Leu Ser His
 5 10 15
 tcc cca ggt acc tcc aat cag agc cag ccc tgt tct cca aag cca atc 393
 Ser Pro Gly Thr Ser Asn Gln Ser Gln Pro Cys Ser Pro Lys Pro Ile
 20 25 30
 cgc ctg gtt cag gac ctc cca gag gag ctg gtg cat gca ggc tgg gag 441
 Arg Leu Val Gln Asp Leu Pro Glu Glu Leu Val His Ala Gly Trp Glu
 35 40 45 50
 aag tgc tgg agc cgg agg gag aat cgt ccc tac tac ttc aac cga ttc 489
 Lys Cys Trp Ser Arg Arg Glu Asn Arg Pro Tyr Tyr Phe Asn Arg Phe
 55 60 65
 acc aac cag tcc ctg tgg gag atg ccc gtg ctg ggg cag cac gat gtg 537
 Thr Asn Gln Ser Leu Trp Glu Met Pro Val Leu Gly Gln His Asp Val
 70 75 80
 att tcg gac cct ttg ggg ctg aat gcg acc cca ctg ccc caa gac tca 585
 Ile Ser Asp Pro Leu Gly Leu Asn Ala Thr Pro Leu Pro Gln Asp Ser
 85 90 95
 agc ttg gtg gaa act ccc ccg gct gag aac aag ccc aga aag cgg cag 633
 Ser Leu Val Glu Thr Pro Pro Ala Glu Asn Lys Pro Arg Lys Arg Gln
 100 105 110
 ctc tcg gaa gag cag cca agc ggc aat ggt gtg aag aag ccc aag att 681
 Leu Ser Glu Glu Gln Pro Ser Gly Asn Gly Val Lys Lys Pro Lys Ile
 115 120 125 130
 gaa atc cca gtg aca ccc aca ggc cag tcg gtg ccc agc tcc ccc agt 729
 Glu Ile Pro Val Thr Pro Thr Gly Gln Ser Val Pro Ser Ser Pro Ser
 135 140 145
 atc cca gga acc cca acg ctg aag atg tgg ggt acg tcc cct gaa gat 777
 Ile Pro Gly Thr Pro Thr Leu Lys Met Trp Gly Thr Ser Pro Glu Asp
 150 155 160
 aaa cag cag gca gct ctc cta cga ccc act gag gtc tac tgg gac ctg 825
 Lys Gln Gln Ala Ala Leu Leu Arg Pro Thr Glu Val Tyr Trp Asp Leu
 165 170 175
 gac atc cag acc aat gct gtc atc aag cac cgg ggg cct tca gag gtg 873
 Asp Ile Gln Thr Asn Ala Val Ile Lys His Arg Gly Pro Ser Glu Val
 180 185 190
 ctg ccc ccg cat ccc gaa gtg gaa ctg ctc cgc tct cag ctc atc ctg 921
 Leu Pro Pro His Pro Glu Val Glu Leu Leu Arg Ser Gln Leu Ile Leu
 195 200 205 210
 aag ctt cgg cag cac tat cgg gag ctg tgc cag cag cga gag ggc att 969
 Lys Leu Arg Gln His Tyr Arg Glu Leu Cys Gln Gln Arg Glu Gly Ile
 215 220 225
 gag cct cca cgg gag tct ttc aac cgc tgg atg ctg gag cgc aag gtg 1017
 Glu Pro Pro Arg Glu Ser Phe Asn Arg Trp Met Leu Glu Arg Lys Val
 230 235 240
 gta gac aaa gga tct gac ccc ctg ttg ccc agc aac tgt gaa cca gtc 1065
 Val Asp Lys Gly Ser Asp Pro Leu Leu Pro Ser Asn Cys Glu Pro Val
 245 250 255
 gtg tca cct tcc atg ttt cgt gaa atc atg aac gac att cct atc agg 1113
 Val Ser Pro Ser Met Phe Arg Glu Ile Met Asn Asp Ile Pro Ile Arg
 260 265 270

tta tcc cga atc aag ttc cgg gag gaa gcc aag cgc ctg ctc ttt aaa 1161
 Leu Ser Arg Ile Lys Phe Arg Glu Glu Ala Lys Arg Leu Leu Phe Lys
 275 280 285 290
 tat gcg gag gcc gcc agg cgg ctc atc gag tcc agg agt gca tcc cct 1209
 Tyr Ala Glu Ala Ala Arg Arg Leu Ile Glu Ser Arg Ser Ala Ser Pro
 295 300 305
 gac agt agg aag gtg gtc aaa tgg aat gtg gaa gac acc ttt agc tgg 1257
 Asp Ser Arg Lys Val Val Lys Trp Asn Val Glu Asp Thr Phe Ser Trp
 310 315 320
 ctt cgg aag gac cac tca gcc tcc aag gag gac tac atg gat cgc ctg 1305
 Leu Arg Lys Asp His Ser Ala Ser Lys Glu Asp Tyr Met Asp Arg Leu
 325 330 335
 gag cat ctg cgg agg cag tgt ggc ccc cac gtc tcg gcc gca gcc aag 1353
 Glu His Leu Arg Arg Gln Cys Gly Pro His Val Ser Ala Ala Ala Lys
 340 345 350
 gac tcc gtg gaa ggc atc tgc agt aag atc tac cac atc tcc ctg gag 1401
 Asp Ser Val Glu Gly Ile Cys Ser Lys Ile Tyr His Ile Ser Leu Glu
 355 360 365 370
 tac gtc aaa cgg atc cga gag aag cac ctt gcc atc ctc aag gaa aac 1449
 Tyr Val Lys Arg Ile Arg Glu Lys His Leu Ala Ile Leu Lys Glu Asn
 375 380 385
 aac atc tca gag gag gtg gag gcc cct gag gtg gag ccc cgc cta gtg 1497
 Asn Ile Ser Glu Glu Val Glu Ala Pro Glu Val Glu Pro Arg Leu Val
 390 395 400
 tac tgc tac cca gtc cgg ctg gct gtg tct gca ccg ccc atg ccc agc 1545
 Tyr Cys Tyr Pro Val Arg Leu Ala Val Ser Ala Pro Pro Met Pro Ser
 405 410 415
 gtg gag atg cac atg gag aac aac gtg gtc tgc atc cgg tat aag gga 1593
 Val Glu Met His Met Glu Asn Asn Val Val Cys Ile Arg Tyr Lys Gly
 420 425 430
 gag atg gtc aag gtc agc cgc aac tac ttc agc aag ctg tgg ctc ctt 1641
 Glu Met Val Lys Val Ser Arg Asn Tyr Phe Ser Lys Leu Trp Leu Leu
 435 440 445 450
 tac cgc tac agc tgc att gat gac tct gcc ttt gag agg ttc ctg ccc 1689
 Tyr Arg Tyr Ser Cys Ile Asp Asp Ser Ala Phe Glu Arg Phe Leu Pro
 455 460 465
 cgg gtc tgg tgt ctt ctc cga cgg tac cag atg atg ttc ggc gtg ggc 1737
 Arg Val Trp Cys Leu Leu Arg Arg Tyr Gln Met Met Phe Gly Val Gly
 470 475 480
 ctc tac gag ggg act ggc ctg cag gga tcg ctg cct gtg cat gtc ttt 1785
 Leu Tyr Glu Gly Thr Gly Leu Gln Gly Ser Leu Pro Val His Val Phe
 485 490 495
 gag gcc ctc cac cga ctc ttt ggc gtc agc ttc gag tgc ttc gcc tca 1833
 Glu Ala Leu His Arg Leu Phe Gly Val Ser Phe Glu Cys Phe Ala Ser
 500 505 510
 ccc ctc aac tgc tac ttc cgc cag tac tgt tct gcc ttc ccc gac aca 1881
 Pro Leu Asn Cys Tyr Phe Arg Gln Tyr Cys Ser Ala Phe Pro Asp Thr
 515 520 525 530
 gac ggc tac ttt ggc tcc cgc ggg ccc tgc cta gac ttt gct cca ctg 1929
 Asp Gly Tyr Phe Gly Ser Arg Gly Pro Cys Leu Asp Phe Ala Pro Leu
 535 540 545 550
 agt ggt tca ttt gag gcc aac cct ccc ttc tgc gag gag ctc atg gat 1977
 Ser Gly Ser Phe Glu Ala Asn Pro Pro Phe Cys Glu Glu Leu Met Asp
 550 555 560
 gcc atg gtc tct cac ttt gag aga ctg ctt gag agc tca ccg gag ccc 2025
 Ala Met Val Ser His Phe Glu Arg Leu Leu Glu Ser Ser Pro Glu Pro
 565 570 575

ctg tcc ttc atc gtg ttc atc cct gag tgg cgg gaa ccc cca aca cca 2073
 Leu Ser Phe Ile Val Phe Ile Pro Glu Trp Arg Glu Pro Pro Thr Pro
 580 585 590
 gcg ctc acc cgc atg gag cag agc cgc ttc aaa cgc cac cag ttg atc 2121
 Ala Leu Thr Arg Met Glu Gln Ser Arg Phe Lys Arg His Gln Leu Ile
 595 600 605 610
 ctg cct gcc ttt gag cat gag tac cgc agt ggc tcc cag cac atc tgc 2169
 Leu Pro Ala Phe Glu His Glu Tyr Arg Ser Gly Ser Gln His Ile Cys
 615 620 625
 aag aag gag gaa atg cac tac aag gcc gtc cac aac acg gct gtg ctc 2217
 Lys Lys Glu Glu Met His Tyr Lys Ala Val His Asn Thr Ala Val Leu
 630 635 640
 ttc cta cag aac gac cct ggc ttt gcc aag tgg gcg ccg acg cct gaa 2265
 Phe Leu Gln Asn Asp Pro Gly Phe Ala Lys Trp Ala Pro Thr Pro Glu
 645 650 655
 cgg ctg cag gag ctg agt gct gcc tac cgg cag tca ggc cgc agc cac 2313
 Arg Leu Gln Glu Leu Ser Ala Ala Tyr Arg Gln Ser Gly Arg Ser His
 660 665 670
 agc tct ggt tct tcc tca tgc tcc tcc tgc gag gcc aag gac cgg gac 2361
 Ser Ser Gly Ser Ser Ser Ser Ser Ser Ser Ser Glu Ala Lys Asp Arg Asp
 675 680 685 690
 tgc ggc cgt gag cag ggt cct agc cgc gag cct cac ccc act taa 2406
 Ser Gly Arg Glu Gln Gly Pro Ser Arg Glu Pro His Pro Thr
 695 700
 catatcctgc ggggaggagg agccccagggt gtgctagtct ggactgctgg gactcgggcc 2466
 cctggggcct cagagggacc ccggctgcca ctgacatatg aagattatgg ttctgccagg 2526
 gctccccctcc ctgcctgtcc ccaagtcctc acctcaaact ccctccaagt cccatgtata 2586
 taggtcctga tgccttccca accccgcccc tcaccctgtt gccaccttgt ttcatttgta 2646
 aaaggaaata cagaaacccc ccc 2669

<210> 4
 <211> 26
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthesized oligonucleotide

<400> 4
 ccgaattcat ggccaatgag aatcac 26

<210> 5
 <211> 26
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthesized oligonucleotide

<400> 5
 ccgtcgactt aagtgggggtg aggctc 26

<210> 6
 <211> 34

<212> DNA

<213> Artificial sequence

<220>

<223> Synthesized oligonucleotide

<400> 6

cgaggatccg ttcaggacct cccagaggac gcta

34

<210> 7

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<223> Synthesized oligonucleotide

<400> 7

cgagaattcc gaaatcacat cgtgctgccc cag

33